THE PARTITIONED REGRESSSION MODEL

Consider taking a regression equation in the form of

$$(1) \qquad y = [\, X_1 \quad X_2 \,] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \varepsilon = X_1\beta_1 + X_2\beta_2 + \varepsilon.$$

Here $[X_1, X_2] = X$ and $[\beta_1', \beta_2']' = \beta$ are obtained by partitioning the matrix X and vector β of the equation $y = X\beta + \varepsilon$ in a conformable manner. The normal equations $X'X\beta = X'y$ can be partitioned likewise. Writing the equations without the surrounding matrix braces gives

$$(2) \qquad X_1'X_1\beta_1 + X_1'X_2\beta_2 = X_1'y,$$

$$(3) \qquad X_2'X_1\beta_1 + X_2'X_2\beta_2 = X_2'y.$$

From (2), we get the equation $X_1'X_1\beta_1 = X_1'(y - X_2\beta_2)$ which gives an expression for the leading subvector of $\hat\beta$:

$$(4) \qquad \hat\beta_1 = (X_1'X_1)^{-1}X_1'(y - X_2\hat\beta_2).$$

To obtain an expression for $\hat\beta_2$, we must eliminate β_1 from equation (3). For this purpose, we multiply equation (2) by $X_2'X_1(X_1'X_1)^{-1}$ to give

$$(5) \qquad X_2'X_1\beta_1 + X_2'X_1(X_1'X_1)^{-1}X_1'X_2\beta_2 = X_2'X_1(X_1'X_1)^{-1}X_1'y.$$

When the latter is taken from equation (3), we get

$$(6) \qquad \left\{ X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2 \right\}\beta_2 = X_2'y - X_2'X_1(X_1'X_1)^{-1}X_1'y.$$

On defining

$$(7) \qquad P_1 = X_1(X_1'X_1)^{-1}X_1',$$

can we rewrite (6) as

$$(8) \qquad \left\{ X_2'(I - P_1)X_2 \right\}\beta_2 = X_2'(I - P_1)y,$$

whence

$$(9) \qquad \hat\beta_2 = \left\{ X_2'(I - P_1)X_2 \right\}^{-1} X_2'(I - P_1)y.$$

Now let us investigate the effect that conditions of orthogonality amongst the regressors have upon the ordinary least-squares estimates of the regression parameters. Consider a partitioned regression model, which can be written as

$$(10) \qquad y = [\, X_1, X_2 \,] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \varepsilon = X_1\beta_1 + X_2\beta_2 + \varepsilon.$$

It can be assumed that the variables in this equation are in deviation form. Imagine that the columns of X_1 are orthogonal to the columns of X_2 such that $X_1'X_2 = 0$. This is the same as assuming that the empirical correlation between variables in X_1 and variables in X_2 is zero.

The effect upon the ordinary least-squares estimator can be seen by examining the partitioned form of the formula $\hat{\beta} = (X'X)^{-1}X'y$. Here we have

$$(11) \qquad X'X = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} [X_1 \quad X_2] = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix} = \begin{bmatrix} X_1'X_1 & 0 \\ 0 & X_2'X_2 \end{bmatrix},$$

where the final equality follows from the condition of orthogonality. The inverse of the partitioned form of $X'X$ in the case of $X_1'X_2 = 0$ is

$$(12) \qquad (X'X)^{-1} = \begin{bmatrix} X_1'X_1 & 0 \\ 0 & X_2'X_2 \end{bmatrix}^{-1} = \begin{bmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & (X_2'X_2)^{-1} \end{bmatrix}.$$

We also have

$$(13) \qquad X'y = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} y = \begin{bmatrix} X_1'y \\ X_2'y \end{bmatrix}.$$

On combining these elements, we find that

$$(14) \qquad \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & (X_2'X_2)^{-1} \end{bmatrix} \begin{bmatrix} X_1'y \\ X_2'y \end{bmatrix} = \begin{bmatrix} (X_1'X_1)^{-1}X_1'y \\ (X_2'X_2)^{-1}X_2'y \end{bmatrix}.$$

In this special case, the coefficients of the regression of y on $X = [X_1, X_2]$ can be obtained from the separate regressions of y on X_1 and y on X_2.

It should be understood that this result does not hold true in general. The general formulae for $\hat{\beta}_1$ and $\hat{\beta}_2$ are those which we have given already under (4) and (9):

$$(15) \qquad \begin{aligned} \hat{\beta}_1 &= (X_1'X_1)^{-1}X_1'(y - X_2\hat{\beta}_2), \\ \hat{\beta}_2 &= \left\{ X_2'(I - P_1)X_2 \right\}^{-1} X_2'(I - P_1)y, \quad P_1 = X_1(X_1'X_1)^{-1}X_1'. \end{aligned}$$

It can be confirmed easily that these formulae do specialise to those under (14) in the case of $X_1'X_2 = 0$.

The purpose of including X_2 in the regression equation when, in fact, interest is confined to the parameters of β_1 is to avoid falsely attributing the explanatory power of the variables of X_2 to those of X_1.

Let us investigate the effects of erroneously excluding X_2 from the regression. In that case, the estimate will be

$$(16) \qquad \begin{aligned} \tilde{\beta}_1 &= (X_1'X_1)^{-1}X_1'y \\ &= (X_1'X_1)^{-1}X_1'(X_1\beta_1 + X_2\beta_2 + \varepsilon) \\ &= \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2 + (X_1'X_1)^{-1}X_1'\varepsilon. \end{aligned}$$

On applying the expectations operator to these equations, we find that

$$(17) \qquad\qquad E(\tilde{\beta}_1) = \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2,$$

since $E\{(X_1'X_1)^{-1}X_1'\varepsilon\} = (X_1'X_1)^{-1}X_1'E(\varepsilon) = 0$. Thus, in general, we have $E(\tilde{\beta}_1) \neq \beta_1$, which is to say that $\tilde{\beta}_1$ is a biased estimator. The only circumstances in which the estimator will be unbiased are when either $X_1'X_2 = 0$ or $\beta_2 = 0$. In other circumstances, the estimator will suffer from a problem which is commonly described as *omitted-variables bias*.

We need to ask whether it matters that the estimated regression parameters are biased. The answer depends upon the use to which we wish to put the estimated regression equation. The issue is whether the equation is to be used simply for predicting the values of the dependent variable y or whether it is to be used for some kind of structural analysis.

If the regression equation purports to describe a structural or a behavioral relationship within the economy, and if some of the explanatory variables on the RHS are destined to become the instruments of an economic policy, then it is important to have unbiased estimators of the associated parameters. For these parameters indicate the leverage of the policy instruments. Examples of such instruments are provided by interest rates, tax rates, exchange rates and the like.

On the other hand, if the estimated regression equation is to be viewed solely as a predictive device—that it to say, if it is simply an estimate of the function $E(y|x_1, \ldots, x_k)$ which specifies the conditional expectation of y given the values of $x_1, \ldots, x_n$—then, provided that the underlying statistical mechanism which has generated these variables is preserved, the question of the unbiasedness the regression estimates does not arise.